



US 20030004652A1

(19) **United States**(12) **Patent Application Publication** (10) Pub. No.: **US 2003/0004652 A1****Brunner et al.**

(43) Pub. Date:

**Jan. 2, 2003**(54) **SYSTEMS AND METHODS FOR  
MONITORING BEHAVIOR INFORMATICS**

(52) U.S. Cl. .... 702/19; 705/2

(76) Inventors: **Daniela Brunner, Bronx, NY (US);  
Vijay Gondhalekar, New York, NY  
(US); Emer Leahy, Bedford, NY (US)**(57) **ABSTRACT**

Correspondence Address:  
**ROPES & GRAY  
ONE INTERNATIONAL PLACE  
BOSTON, MA 02110-2624 (US)**

A system and method used to assess animal behavior includes a module having sensors that collects a variety of physical and biological data from a test subject. Interpretation of the data is provided to assess the test subject's behavior, neurology, biochemistry and physiology. The module is useful in observing the effects of a drug on the test animal and providing information on the drug's signature. Another advantage is module's portability that allows it to be used in standard laboratory cages. (NOT SURE ABOUT THIS PORTABILITY) This portability allows the animal to be tested in its own habitat, that can reduce any erroneous data due to stressing the animal when removed to a test cage. Additionally, the module's design allows for parallel data collection and interpretation from several laboratory animals undergoing different experiments. Multi-dimensional modeling of the test subject based the system's interpretation of the data allows pattern recognition of the drug signature, and predictive drug analysis.

(21) Appl. No.: **10/147,336**(22) Filed: **May 15, 2002****Related U.S. Application Data**

(60) Provisional application No. 60/291,039, filed on May 15, 2001. Provisional application No. 60/326,271, filed on Oct. 1, 2001.

**Publication Classification**(51) Int. Cl.<sup>7</sup> ..... **G06F 17/60; G06F 19/00;  
G01N 33/48; G01N 33/50**

PGPUB-DOCUMENT-NUMBER: 20030004652

PGPUB-FILING-TYPE: new

DOCUMENT-IDENTIFIER: US 20030004652 A1

TITLE: Systems and methods for monitoring behavior informatics

PUBLICATION-DATE: January 2, 2003

INVENTOR-INFORMATION:

NAME	CITY	STATE	COUNTRY
RULE-47			
Brunner, Daniela	Bronx	NY	US
Gondhalekar, Vijay	New York	NY	US
Leahy, Emer	Bedford	NY	US

US-CL-CURRENT: 702/19, 705/2

ABSTRACT:

A system and method used to assess animal behavior includes a module having sensors that collects a variety of physical and biological data from a test subject. Interpretation of the data is provided to assess the test subject's behavior, neurology, biochemistry and physiology. The module is useful in observing the effects of a drug on the test animal and providing information on the drug's signature. Another advantage is module's portability that allows it to be used in standard laboratory cages. (NOT SURE ABOUT THIS PORTABILITY) This portability allows the animal to be tested in its own habitat, that can reduce any erroneous data due to stressing the animal when removed to a test cage. Additionally, the module's design allows for parallel data collection and interpretation from several laboratory animals undergoing different experiments. Multi-dimensional modeling of the test subject based the system's interpretation of the data allows pattern recognition of the drug signature, and predictive drug analysis.

----- KWIC -----

Detail Description Paragraph - DETX (29):

[0129] The apparatus and method used to assess animal behavior includes a module having one or more sensors that collectively obtain a variety of physical and biological data from a test subject. The sensors of the apparatus also may provide for the ability to simultaneously obtaining data pertaining to different states in the data collecting hardware. Such states include but are not limited to the position of different mechanical parts of the hardware, active video pixels that may or may not correlate with the subjects position and movement, the signals arising from telemetric devices such as cardiac signals and any other analogical or digital signal that can be collected simultaneously (or within a close time window). Time stamped data is stored in a secondary database (where primary database refer to a putative off line storage of video images and other raw data) that allows sophisticated mining. Complex signals are summarized in probability maps using, for example, hidden Markov models (HMMs) to assign transition probabilities between different combinations of the many variables collected. Preferably low-probability states are not dropped from the model, as these low-probability states will help define the more subtle characteristics of the drug or gene signature. These probability maps will be differential, a result of the comparison between

a control group and the experimental group. Multiple-dimensional modeling of the test subject based on the system's interpretation of the data allows pattern recognition of the drug signature, predictive drug analysis, and interpretation of the phenotype of a genetically engineered animal.

Detail Description Paragraph - DETX (38):

[0138] A "clustering algorithm" finds groups of items that are similar. For example, clustering could be used to group drugs according to effect on ambulatory movements, pain sensitivity, convulsions, cardiac output and QT interval. It divides a data set so that records with similar content are in the same group, and groups are as different as possible from each other. When the categories are unspecified, this is sometimes referred to as unsupervised clustering. When the categories are specified a priori, this is sometimes referred to as supervised clustering.

Detail Description Paragraph - DETX (130):

[0230] Supervised learning requires the buildup of a library of readily classified data sets for input into the neural net. Although more economic in terms of the amount of data needed, supervised learning implies that only pre-determined classes can be ascribed to unseen data. To allow for the possibility of finding a novel therapeutic class, such as "antidepressant drugs with anti-manic component" unsupervised clustering could be more appropriate.

Detail Description Paragraph - DETX (133):

[0233] In still other embodiments, the subject method utilizes K-means and fuzzy clustering. Gaussian mixture models are a common version of this. These techniques are "unsupervised" clustering methods. They assume the user has no outputs, but would like to group the data anyway according to inputs that are similar to each other. The idea is to choose a model for each cluster. For example, each cluster may consist of points inside a hyper-sphere centered at some location in the input space. These methods automatically determine the number of clusters, place them in the correct places, and determine which points belong to which clusters. An advantage to these techniques is that they can be efficient algorithms and can do a good job of finding clusters. This is a method of choice when the user does not have a priori information about the classes

Detail Description Paragraph - DETX (134):

[0234] Another embodiment utilizes the hierarchical clustering Serial Linkage Method. This is an unsupervised clustering method in the same sense as K-means and fuzzy clustering. Here individual points are joined to each other by being close to each other in the input space. As these points are joined together, they define clusters. As the algorithm continues, the clusters are joined together to form larger clusters. Compared to K-means and fuzzy clustering, hierarchical clustering has the advantage that clusters can have arbitrary non-predefined shapes and the result correctly shows "clusters of clusters." A disadvantage to these methods is they tend to be more sensitive to noise.

Detail Description Paragraph - DETX (278):

[0370] The illustrated Control Computer system 1402 also includes a Data Collection module 1436. Data from the Computer Vision modules and Low Level Instrument interface 1432 can be passed to the Data Collection module 1436, which can serve as a management system for system data. Other modules which require sensor readings or computer vision output can also connect to the Data Collection Module 1436, which synchronizes the data from its various inputs, and acts as a storage Raw Data and Sequence Database 1436, by storing experimental data from each individual subject for further processing.

Detail Description Paragraph - DETX (279):

[0371] The data corresponding to reference compounds stored in the Raw Data and Sequence Database 1438 can be queried by a Supervised Machine Learning module 1444. Supervised learning techniques, such as Bayes Classifiers, Support Vector Machines, and Mixture models can be used to extract those characteristics of the experimental data that permits successful classification into pre-determined therapeutic classes (e.g. antidepressants). The summarized data for each drug, the drug signature, is stored in the Drug Class Signature Database 1448. The data corresponding to a test or novel compound stored in the Raw Data and Sequence Database 1436 can be queried by a Module 1447 that extracts the appropriate summary data or novel or test drug signature. Such signature is compared against the reference Drug Class Signature Database 1448, by a Machine Learning module 1440, which may used diverse learning algorithms, and which classifies the novel or test drug signature into the existing drug classes previously defined by the Supervised Machine Learning module 1444.